AFRL-RH-WP-TR-2014-0123

A COMPUTATIONAL DUAL-PROCESS MODEL OF SOCIAL INTERACTION

Stephen Deutsch
BBN Technologies
10 Moulton Street
Cambridge, MA 02138

Michael Young, Ph.D.
Human-Centered ISR Division
Human Trust and Interaction Branch
Wright-Patterson AFB OH 45433

JANUARY 2014

INTERIM REPORT
January 2012 through January 2014

AIR FORCE RESEARCH LABORATORY
711th HUMAN PERFORMANCE WING
HUMAN EFFECTIVENESS DIRECTORATE
AIR FORCE MATERIEL COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
UNITED STATES AIR FORCE

Notice and Signature Page

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

Qualified requestors may obtain copies of this report from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RH-WP-TR-2014-0123 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//signature//	//signature//
MICHAEL YOUNG, Ph.D. Work Unit Manager Human Trust and Interaction Branch	LOUISE CARTER, Ph.D. Chief, Human-Centered ISR Division Human Effectiveness Directorate 711th Human Performance Wing Air Force Research Laboratory

This report is published in the interest of scientific and technical information exchange and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. REPORT DATE *(DD-MM-YY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
30 January 2014	Interim	01 January 2012 – 01 January 2014

4. TITLE AND SUBTITLE
A Computational Dual-Process Model of Social Interaction

5a. CONTRACT NUMBER
FA8650-08-D-6801/0020

5b. GRANT NUMBER NA

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

[1]Stephen Deutsch
[2]Michael Young, Ph.D.

5d. PROJECT NUMBER

5e. TASK NUMBER
TO-20

5f. WORK UNIT NUMBER
H06Z (7184X23W)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
[1]BBN Technologies
10 Moulton Street
Cambridge, MA 02138

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
[2]Air Force Materiel Command,
Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Human-Centered ISR Division
Human Trust and Interaction Branch
Wright-Patterson AFB OH 45433

10. SPONSORING/MONITORING AGENCY ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RH-WP-TR-2014-0123

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A. Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
88ABW-2014-5663; Cleared 3 December 2014

14. ABSTRACT (Maximum 200 words)
Dual-process models postulate two distinct modes of information processing, one automatically invoked, highly parallel, and not under conscious control, and the other consciously invoked and limited to serial processing. Stereotype-driven behaviors have been identified as an example of an implicitly activated and automatically invoked behavior. In this effort, we describe the modeling of stereotypes and prejudice, principally with respect to race and religion, in social interactions. We developed a computational model of these implicitly activated behaviors and how they are sometimes refined, even overridden, by often concurrent, explicitly-driven processes. The test-bed for the work reported on here is the Operator Model Architecture (OMAR), a simulation framework for agent-based modeling. OMAR was used to facilitate the building of the computational models of agents, visualized as avatars, which pursue goals that drive their behaviors in social interactions. In carrying out their proactive agendas the OMAR agents respond to perceptual stimuli that activate implicit attitudes which may, in turn, be overturned by stimuli that activate explicit attitudes leading to alternate behaviors. Our goal has been to model both the proactive and reactive processes, and their concurrent implicit and explicit components, in social interactions from perceptual input through behavioral consequences.

15. SUBJECT TERMS
Operator Model Architecture (OMAR), human social cultural behavioral modeling and simulation, cognitive models, tacit knowledge

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Michael Young
U	U	U	SAR	26	**19b. TELEPHONE NUMBER** *(Include Area Code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

THIS PAGE INTENTIONALLY LEFT BLANK.

TABLE OF CONTENTS

1.0 INTRODUCTION

Dual-process models postulate two distinct modes of information processing, one automatically invoked, highly parallel, and not under conscious control, and the other consciously invoked and limited to serial processing. The roots of these models can be traced back to the earlier research on automaticity in experimental psychology (Logan, 1988; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Following these initial research efforts in experimental psychology, the concept of automaticity then had a broader impact on research in social psychology (Bargh & Chartrand, 1999). In particular, stereotype-driven behaviors were identified as implicitly activated and automatically invoked (Devine, 1989; Fiske, 1998). It is the modeling of stereotypes and prejudice, principally with respect to race and religion, in social interactions that is the work reported on here. We have been concerned with developing a *computational* model of these implicitly activated behaviors and how they are sometimes refined, even overridden, by often concurrent, explicitly-driven processes.

The test-bed for the work reported on here is the Operator Model Architecture (OMAR), a simulation framework for agent-based modeling that has been used for many human performance modeling research efforts (Deutsch, 1998; Deutsch & Pew, 2002) in the past. It is a system that we have developed, refined, and employed for over 20 years. OMAR was used to facilitate the building of the computational models in which the agents, visualized as avatars, pursue the goals that drive their behaviors in social interactions. In carrying out their proactive agendas they respond to perceptual stimuli that activate implicit attitudes driving their behaviors that may, in turn, be overturned by stimuli that activate explicit attitudes leading to alternate behaviors. Our aim has been to model the proactive and reactive processes and their concurrent implicit and explicit components in social interactions from perceptual input through behavioral consequences (Strack & Deutsch, 2004).

Considerable research in social psychology has pursued the development of both single- and dual-process theories and models of social interactions. Each addresses the implicit and explicit components that drive behaviors that we have focused on in our modeling effort. We look first at factors that characterize the processes of implicitly- and explicitly-activated associations leading to action selection. Then we examine the details of the competition among implicitly activated possible outcomes. From there we move on to examine the influence of explicitly processed possible outcomes and the competition between those and the selected implicitly selected outcome. As we will see, each of these particular areas has spawned competing approaches to their interpretation and understanding that has required careful attention as the computational model building was pursued.

We then examine the choices that we have made as seen in how each of these human resources and capabilities are realized in the computational model. With the best effort at theoretical choices in place, we then examine the scenarios that have been developed to demonstrate how the influences of stereotypes drive behaviors and how those behaviors might be influenced by explicitly processed attitudes that may operate concurrently.

In contrast to how individuals first interact with strangers, particularly those who are members of an out-group, interactions with acquaintances, and those identified as in-group strangers may be quite different. We examine how we model the acquaintanceships that individuals maintain with one another; relationships that persist over an extended period of time, often across diverse social situations.

Lastly, we provide an overview of the visualization of the scenarios' human performance models as avatars that portray the social interactions of the individuals involved.

2.0 DUAL-PROCESS THEORIES AND MODELS OF SOCIAL INTERACTION

There is an extensive literature in social psychology on single- and dual-process theories and models of the process from perceptual input to behavioral consequences. However, at this stage in theory and model development there is only limited agreement at a surface level on how the processes might be differentiated and defined. At a more detailed level, there is a significant range of thought in theory and model development that has been derived from a very extensive range of human subject experiments across a number of closely related social science research areas. As is to be expected, the several theories are in flux, with recent arguments focusing on whether there are two distinct systems or whether there are two different types of processes (Evans & Stanovich, 2013). In laying out the theoretical foundation for our computational model we made a best effort to navigate this thicket in several stages that we now outline.

We started from a dual-process perspective and begin by outlining our choice to identify the two processes as implicit and explicit processes: those that proceed unattended and those that are consciously attended. We then focus first on the activation of implicitly activated associations between attitude and behavior and the competition among the possible automatically choices that must be resolved. We then outline our approach to weaving in often concurrently activated explicit behavioral choices and highlight the top-level considerations that drive the design of the computational model. We conclude this section with a discussion of how the behavior that is actually pursued is derived from the competing implicit and explicit processes that lead to the array of possible choices.

2.1 Implicit and Explicit Factors in Action Selection

There has been an extended discussion in the social psychology literature of the principal factors that characterize the two systems or processes in the determination of behaviors (cf., Evans & Stanovich, 2013). And indeed, Kruglanski and Gigerenzer (2011) propose a *single* system, "a unified theoretical approach that explains both intuitive and deliberative judgments as rule based, as opposed to the dual-systems approach of qualitatively different processes." With respect to two-system theories, we note that others (e.g., Gawronski & Bodenhausen, 2011) have cited a variety of other factors as being the essential attributes necessary to properly distinguish and characterize the two distinct processes. Quite recently, Gawronski and Creighton (2013) in introducing dual process theories suggested that "the defining characteristic of these theories is that they divide the mental processes underlying social judgments and behavior into two general categories depending on whether they operate automatically or in a controlled fashion."

Early on, Logan (1988) characterized automatic processing as fast, effortless, autonomous, stereotypic, and unavailable to conscious awareness. This characterization is very much in the spirit of Lieberman's (2009) more recently defined X-system and C-system where he contrasts non-reflective and reflective processes as:

- Parallel processing vs. serial processing
- Fast operating vs. slow operating
- Unaffected by cognitive load vs. altered by cognitive load
- Implicit learning of associations vs. explicit learning of rules
- Pattern matching and pattern completion vs. symbolic logic and propositional reasoning

We determined this characterization a reasonable basis from which to begin our computational modeling effort.

2.2 The Activation of Competing Implicit Associations

The associations that we are concerned with are those between attitudes and behaviors; the responses to input stimuli that trigger attitudes that influence behavioral choices. While the definition of an *attitude* has itself been the subject of considerable attention (c.f., *Social Cognition* (2007, Vol. 25[5]), "What's an Attitude?"), for our purposes an attitude is a specialization of a concept. A concept is a set of features associated with an object. An attitude is a set of beliefs associated with an object. The activation of an attitude is similar to the activation of a concept in that it is a set of associations. However, it differs from activation of a concept in that there is a significant affective component associate with the "features," that is, you normally have an emotionally toned reaction of either liking or disliking the object. In addition, attitudinal associations are not usually semantically based, but instead they are episodically based. They seem to be derived from previous experiences with exemplars from a specific class of objects. A vast variety of things can trigger attitudes from views about politics to views about automation aiding, for example. In the research reported here we are concerned with attitudes concerning ethnic groups. If a specific individual—of any ethnicity—is not perceived as an individual, but rather characterized or associated with a set of group level attributes—then we say they have been perceived in a stereotypic manner

We are concerned with two types of attitudes in this effort. First there are *implicit* attitudes that become automatically activated through repeated exposures in a variety of contexts or situations. Individuals seem to learn these (predominantly negative) associations concerning specific ethnic groups early in their development before they can critically evaluate the validity of the underlying stereotype. (Fiske, 1998). Second, there are *explicit* attitudes that seem to be learned as a result of negative feedback from peers or others in social situations that cause an individual to reevaluate their attitudes towards individuals of specific ethnic groups.

We begin by looking at the automatic response to a set of input stimuli that takes the form of implicit processing. The implicit processing allows for a set of input stimuli, each representing or coding an attribute of the external stimulus or object (e.g., color of skin, or ethnic style of dress), to activate one or more associations each with a potential behavioral outcome and hence, triggering a competition leading to a single selected outcome. The computation of the activation of implicit associations and the competition between them is based on Young's (1998) thesis work which simulates the process through which neuron ensembles compete for dominance. As implemented in the model, the stimuli driving the activation are made up of one or more bit vectors each representing the value for a particular stimulus attribute. The individual attributes of the stimuli may be independently weighted. In this way a particular attribute may have more impact in one situation than another. In Young's original formulation, the individual attributes were grouped in a single bit-vector. The aggregate bit-vector has now been split into factor-specific bit-vectors to enable the individual weighting of the contributing factors.

For the automatic response to a set of stimuli there is typically a small set of possible outcomes. The possible outcomes each have an associated canonical bit-vector representing the set of attitude attribute-values with which each outcome is associated. An outcome is then selected based on the match of the stimuli to the associated canonical bit-vector. Repeated stimuli incrementally activate the set of processes competing to determine behavior. The bit-by-bit

match between the stimuli and the canonical bit-vector from each possible outcome is matched by counting the matching bits; this allows partial matches to produce some activation. The matched bits are counted across the bit-vectors; the weightings are applied to each match count, and then normalized by the total weighted bit count. This determines the additional level of activation for each competing outcome on each round of stimuli. The outcome with its match first crossing its activation threshold is selected and other competitors are inhibited.

The successful competitor among the implicitly driven contenders may simply return a value to be used in determining further actions or directly initiate a procedure with associated attribute-value pairs that drive the response to the input stimuli. Whether or not this implicitly selected response succeeds in determining subsequent behavior is contingent on whether or not a possibly concurrent explicit response selection successfully contests the behavior selection.

2.3 Competing Implicitly- and Explicitly-Selected Behaviors

Confronted with a small array of stimuli, there is a good consensus that there can be a mix of implicitly- and explicitly-activated actions that might then contend to actually manage subsequent behaviors. Unfortunately, there is limited evidence on the details of the process of selecting among the several behavioral choices.

- Might the activation of implicit and explicit processes get underway simultaneously or might one or the other start the competitive process?
- What might the timings of the respective processes be?
- If an implicitly activated response is about to be launched, might its "rightness" be subject to an explicit review and possible override by an explicitly pursued process?
- Is an active critic of an auto-initiated procedure required or is it possible that concurrently activated implicit and explicit threads simply vie to control behavior?
- More specifically, is an explicit critique required to forestall the execution of an implicitly initiated process or might a straightforwardly activated explicit process subsume or simply replace the implicitly chosen process?

Here we might take a moment to attend to the various approaches proposed by researchers to describe the two processes in dual process models. We are interested in how they are characterized and need to examine what can be learned about how the two processes interact to determine behavior. Evans and Stanovich (2013), now speak of Type 1 and Type 2 processing where they had previously spoken of System 1 and System 2. More specifically, they point out that there may indeed be more than two systems supporting the two types of processing. They further emphasize that "Type 2 processing is distinguished from autonomous Type 1 processing by its nature—involving cognitive decoupling and hypothetical thinking—and by its strong loading on the working memory resources that this requires." (Evans & Stanovich, 2013)

Gawronski and Bodenhausen (2006) argue that implicit and explicit attitudes should be understood in terms of their underlying processes which they identify as *associative* and *propositional* processing respectively. Gawronski and Bodenhausen (2011) further argue that "the most important feature that distinguishes between associative and propositional processes is their (in)dependency of subjective truth or falsity."

Lieberman (2009) identifies a neurophysiological-based X-system and C-system as non-reflective and reflective respectively. His characterization of his X-system as non-reflective, fast in processing with slow and implicit learning of associations is very close to Logan's (1988) earlier characterization of automatic processing as fast, effortless, autonomous, stereotypic, and unavailable to conscious awareness. Lieberman's (2009) C-system is reflective, serial in its processing, fast learning with the explicit learning of rules and the use of symbolic logic and propositional reasoning.

In any given instance of selecting a behavior to pursue, while multiple implicit processes may be activated in parallel, two factors constrain the role of explicit processing. First, the subject must have available cognitive capacity to address the issues at hand; second, the subject must have the motivation to pursue an explicit selection (Hardin & Banaji, 2013). With respect to motivation, it may be internally driven of one's own volition or it may be externally driven, that is, by the subjects concerns for what others might think of one's behavior (Plant & Devine, 1998).

2.4 Top-Level Considerations for a Computational Model of Social Interactions

The literature on dual-process theory and models is diverse in its wealth of approaches explored leaving one with much latitude in determining a best approach to constructing a *computational* model. As noted earlier, there have also been threads in the social psychology literature pursuing single process theories and models (e.g., Kruglanski & Gigerenzer, 2011). The computational model requires a top-level framework to accommodate the implicit and explicit processes that come into play as a person's stereotypic attitudes are activated and then maybe mediated through possibly concurrent explicit processes. We have chosen to model the implicit processing as associational within a broader hierarchy of goals and procedures where active goals express an agent's intent and procedures express an agent's behaviors toward accomplishing those goals. A description of how the core associational process is operationalized was provided above.

The explicit system is similarly implemented within the same framework of goals and procedures. Decisions may sometimes be expressed through rules or more simply through the evaluation of short and often simple chains of predicates. At other times decisions may be represented by employing previously acquired heuristics. Our implementation of explicit processing is significantly more "relaxed" than Gawronski and Bodenhausen's (2011) more formal propositional reasoning with a focus on the determination of subjective truth or falsity. They do allow however that "logical consistency is intended to refer more broadly to subjective consistency resulting from any kind of inferential rule that a person considers valid, rather than to strict logical consistency in terms of normative syllogistic rules" (Gawronski & Bodenhausen, 2011).

Broadly speaking, we start from a base of mind as process (Edelman, 1987; 1989) most closely following the early work of Logan (1988) and what we see as the elaboration of that work as pursued in the new area of cognitive social neuroscience by Lieberman (2009; 2010). In our model, there may be multiple implicit selection processes executing concurrently with an explicit selection process that is a conscious, attended process. The explicit process is represented by the execution of a *rule set* that selects from among a small set of explicitly considered behavioral choices. With the selection of approaches to modeling implicit and explicit behaviors, we return to the significant issue of how concurrent implicit and explicit processes might interact to determine behavior.

2.5 Theorizing on Possibly Conflicting Implicitly- and Explicitly-Selected Behaviors

For the present, the focus of our computational modeling effort has been restricted to role of stereotypes in the expression of prejudice where these are seen as implicit association-selected actions that may possibly be mediated through explicitly selected process intervention. In contrast to potentially more cognitively intense areas of decision-making or persuasion where additional considerations such as cognitive elaboration might play a larger role, the modeling of stereotype response allows us to begin the modeling of a somewhat less complex process

Morsella (2005) broadly captures the overall process in his proposed Supramodular Interaction Theory (SIT) where supramodules are "information-processing structures composed of multiple modules and defined in terms of their concerns rather than in terms of their sensory afference." His principle of Parallel Responses Into Skeletal Muscle (PRISM) proposes that phenomenal states cull simultaneously activated response tendencies to yield a single, adaptive skeletomotor action, where the phenomenal states are the product of a mix of conscious and unconscious processes. "Logistically, phenomenal states could be considered as one of the mechanisms solving the problem of integrating processes in a largely parallel brain that must satisfy the demands of a skeletomotor system that can often express actions and goals only one at a time." (Morsella, 2005)

While there are many possible nuances that need consideration in understanding how this interaction might go forward, processes defined as *parallel-competitive* and *default-interventionist* (cf., Evans & Stanovich, 2013) do broadly capture the nature of the competing ideas on how conflicts might be addressed. Evans & Stanovich (2013) define *default-interventionist* theories as assuming that fast Type 1 processing generate intuitive default responses on which subsequent reflective Type 2 processing may or may not intervene. They define *parallel-competitive* theories as assuming that Type1 and 2 processing proceed in parallel, each having their say with conflict resolved if necessary. The distinctions between the two theories appear somewhat subtle and seem not particularly well specified to the extent that one would like as the basis for a computational modeling effort. The locus of the distinction appears to be on the point at which the explicit process might be activated vis-à-vis the implicit processes and the nature and degree of explicit processing that might be necessary to override rapid implicit behavior selection; does the explicit process take on the role of conflict detector and critic or might its operation as the focused conscious process simply override the implicitly selected action with its own explicitly selected action. If the explicit process has the time to complete its selection, it might reasonably be expected to pursue that selected outcome.

We have implemented a process much like the parallel-competitive process, but are open to a model that implements explicit conflict detection and resolution at a future time. That "the dorsal Anterior Cingulate Cortex ACC is typically thought to serve a conflict detection function indicating the need for self-control" (Lieberman, 2010) argues for the future investigation of explicit conflict detection and mediation.

For the present, we argue for concurrent activation of implicit and explicit processes and the pursuit of the explicitly selected outcome when present by the simple expedient of being the process with conscious control. Consistent with this approach, Logan (1988) notes that "the subsequent decision process can be inhibited before it results in an overt response," but that "(a)utomatic processing may be a little harder to control than algorithm-based processing, but only because it tends to be faster and allows less time for an act of control to take effect." Further

supporting parallel-competitive approach, "… these regions appear to be in competition such that if one region is relatively active during task performance, the other tends to be correspondingly deactivated." (Lieberman, 2010) These regions, the brain regions supporting implicit and explicit processing, by deferring to one another suggest that our initial processing model is one of the possible modes of conflict resolution.

Based on these considerations, we have implemented concurrent processes in which there is a competition among implicitly selectable outcomes and among one or more explicitly selectable outcomes from which the default is to follow the explicitly selected outcome when present. Occam's razor further argues for the simplicity of this form of conflict resolution and outcome selection. It will be interesting to find the point at which this parsimonious approach breaks down. That is, when is it that the straightforward explicit selection of a "better" outcome does not satisfice and it must be augmented by an explicit critique of the surfaced implicit selection leading to a newly selected outcome? Or alternatively, how might the concurrent implicit selection processes more subtly interact with and thereby influence the explicit outcome selection?

By its very nature as being autonomous and effortless (Logan, 1988), it is a short step to propose that implicit processing is omnipresent, always learning from a person's experiences and influencing his or her behaviors. It was clear from the earliest dual-process research in stereotypes and prejudice that implicit or explicit processing operating alone can lead to behavioral outcomes (Devine, 1989). The challenge is to better understand when and how implicit processing surfaces and affects explicit processing.

3.0 THE DUAL-PROCESS COMPUTATIONAL MODEL OF SOCIAL INTERACTION

At this point we turn our attention to the implementation of the human performance models and the scenarios that were developed in which to demonstrate and examine their behaviors. We start with a brief introduction to the highlights of the modeling environment that were essential to the computational modeling effort. We then look at the implementation of the possibly concurrent activation of implicit and explicit attitudes and how the conflicts among the behaviors that they might initiate are resolved. With this background material in place, we introduce the office workplace scenarios in which the model behaviors were explored. The following section examines some of the scenarios that have been implemented. We focus on parameterizations of the implicit and explicit model attitudes to demonstrate how these attitudes, when activated, act to drive alternate model behaviors.

In contrast with this first set of scenarios in which individuals are meeting members of an out-group for the first time, we then examine extensions of these scenarios that focus on ongoing relationships among individuals and more specifically, how they engage across different situations. Our interest here is in how two individuals tailor their interactions in diverse situations.

Finally, we introduce the avatars that portray the model's behaviors when OMAR was adapted to operate first with Neverwinter Night and then with OpenSim.

3.1 Overview of Human Performance Modeling in OMAR

OMAR is composed of a simulator and a set of three representation languages that are used to define agents, objects, and the behaviors of both animate and inanimate simulation agents. Each OMAR representation language is built as an extension of the Lisp (Steele, 1990) programming language. The Simple Frame Language (SFL) is a traditional frame language used to define the objects and agents for a scenario. A graphical editor supports the development and review of SFL concept and role hierarchies where concepts define objects and roles are the basis for the slots that define their attribute values. The Simulation Core (SCORE) is a language of goals and procedures that define and drive agent behaviors. Active goals express an agent's intents and procedures determine the agent's actions. The language provides for processes that can operate in parallel and is enriched by a signal-passing protocol to further coordinate procedure control and execution. A browser is available to examine the structure of SCORE goal and procedure hierarchies and the deployment of the signal passing component of the language.

The simulator in which the SCORE language executes was derived from the Actors (Agha, 1986) research program at Massachusetts Institute of Technology (MIT). The signal-passing component of the language implements a data-flow architecture (Arvind & Culler, 1983). Agents typically have a mix of goals and procedures that express their proactive goals. They also have a mix of goals and procedures that channel their responses to events that impinge on them, a part of which is concerned with channeling their response to perceptual stimuli. This mix of proactive goals and procedures that guide an agent's purposeful actions and the reactive goals that manage an agent's response to events constitutes what the agent knows how to do (Glenberg, 1997). The computational model, as developed in OMAR, is thus grounded in a process driven architecture (Edelman, 1987, 1989) where rule-based and declarative knowledge play a supporting, rather than a controlling role in an agent's proactive and reactive behaviors. FLEX, also built on Lisp, provides a rule-based language for further defining explicit reasoning behaviors.

The OMAR simulator provides an environment for executing scenarios in which the human models, primary simulation agents, play out their roles interacting with one another and the other inanimate agents and objects that complete the simulation environment. Our focus here will be on the SCORE language and how it is used to implement the competition among the activations of possibly concurrent implicitly and explicitly-activated behaviors. There will be examples of the OMAR Task and Event Timelines that provide detailed insight into agent goal and procedure execution and signal-passing operations.

3.2 Modeling the Processes of Implicit and Explicit Activations in OMAR

The perception of an acquaintance or a stranger may activate a mix of attitudes and implicitly and explicitly selected responses driven by those attitudes. We first examine the implicitly processed aspect of the model, then the explicitly processed aspect of the model, and then how any conflict between them is resolved.

The activation of implicit attitude-driven behavioral choices is an associative matching process as described in detail above. It is accomplished through a signal-passing paradigm enabling in-parallel communication across the procedure hierarchy—the implicitly processed attitude stimuli are received by each of the potential choice procedures at each cycle as each seeks to cross its activation threshold. It is a typical association network. As the activation of the first of the associative processes reaches threshold, the selected procedure begins execution. It is a winner-take-all competition. An inhibit signal from the first activated outcome procedure shuts down the competing implicitly-driven activation processes.

Explicit attitude activation is modeled as a rule-based representation of explicit processing. It is typically a slower process than the concurrent implicit processes. Hence, the explicitly selected outcome procedure will be initiated shortly after the implicitly selected outcome procedure is already underway. It typically lags the start of the implicitly selected outcome by a few hundred milliseconds. It is not the activation processes that are in conflict; they execute concurrently. The conflict is among the selected outcomes of the activation procedures.

Whereas the competition among implicitly activated outcomes was settled through an associative matching process, the conflict between implicitly and explicitly selected outcomes is based on the priorities of their respective outcome procedures. Computationally, the implicit and explicit activations of attitudes are modeled as SCORE procedures where the selected outcome procedures are defined as being in conflict. The SCORE default behavior specifies that a lower priority procedure is to pause until the higher priority procedure with which it is in conflict has completed. Using an additional feature of the SCORE language, we can have the lower priority procedure terminate rather than simply defer its execution when there is a conflict to achieve the desired winner-take-all outcome.

When only implicit attitudes are activated it is the associatively selected outcome that defines further agent behavior. When there are contending implicitly and explicitly selected outcomes, it is the explicitly chosen procedure, the one consciously attended to, that thereby controls subsequent behavior.

Through these coding strategies, there is concurrency in the activations of implicit and explicit procedures and resolution of the conflicts that arise; in the first case among competing implicitly selected outcomes and then between that outcome and an explicitly selected outcome (when present).

3.3 Examining Stereotypic Response and its Mediation in an Office Scenario

We have created an office environment with two office workers, Jane and Joe, and two new interns, Naheed and Darnell, who are each coming in for a first day at work. The focus of the scenarios is on the office workers' behaviors with respect to the new interns where their behaviors are determined in part by the activation and interplay of their implicit and explicit attitudes toward the newly arriving interns. Their implicitly selected behaviors are reflective of their stereotypic attitudes that surface as they perceive the new interns for the first time. At the same time they may have explicit attitudes that come into play and possibly override the implicitly selected actions. The several scenarios are differentiated by the particular settings from among these implicit and explicit attitude parameters.

Roles for the two office workers are assigned a to-do list that each office worker queries on arriving at work. Either office worker can be the greeter for the new interns. Each intern can then be assigned to either of the office workers through the to-do list procedure. One of the interns, Darnell, is a Black male and the other, Naheed, is a Muslim female. The office workers each have attitudes towards race and religion, always implicit and optionally explicit as well. As the interns each arrive for work, the office workers perceive attributes of race and religion on their initial encounter. The perception of visual cues activates the related implicit and explicit attitudes. Additionally, the office workers may or may not be sensitive to the attire (business vs. business casual) of the new interns. Depending on how the office worker's implicit and explicit attitudes interact in determining his or her behaviors, the office worker may assign more or less favorable assignments to the intern.

3.4 Scenarios that Demonstrate the Interplay of Implicit and Explicit Attitudes

We begin by taking a look at two scenarios and how agent parameterization impacts behavioral outcome choices. On arriving at work, Jane consults her to-do list for the day's activities. Based on her to-do list, she finds that she is to greet the new interns and is to take charge of both of them. Alternatively, it might have been the case that Joe was charged with greeting the interns and might have been responsible for one or both of them. In the two scenarios that we will explore, Jane will greet and introduce herself to each of the interns as they individually arrive for work. She will then provide each with a work place and briefly discuss their work assignments.

In each scenario, Darnell is the first to arrive and once settled at a workplace is followed shortly after by Naheed's arrival. We are interested in Jane's implicit and explicit attitudes toward the interns' race and religion and how that impacts her choices for task assignments for the interns. In addition to attitudes with respect to race and religion, Jane can be further characterized by her attitude toward the attire of her new interns where the attire may be business-casual or business-formal. In these scenarios, she is sensitive to the attire of a male intern, Darnell, but not a female intern.

Jane's implicit and explicit (when present) attitudes are triggered by her perception of each intern as he or she first appears at the office. It can be the case that Jane has an implicit, but not an explicit attitude toward a particular perceived attribute or only an explicit altitude with respect to an attribute. If Jane has both an explicit and implicit attitude with respect to an attribute, there may or may not be a conflict between the selected outcomes to pursue. That is, the implicit and explicit attitudes may lead to congruent behavioral outcomes. In this first scenario, Jane has an implicit, but not an explicit attitude toward Darnell's attire.

In our model, a person's particular attitude, implicit or explicit, toward another person is expressed as a bit-vector representing the valence with respect to the attitude. A more favorable attitude with respect to an attribute is represented by more on-bits. An attitude represented by more bits will, in effect, have more weighting in the activation process, although, as we have seen above, each attribute may have a further weight applied to account for situation-specific factors.

In the first scenario, Jane is defined as having an implicit and explicit attitude toward Blacks, each with a value of four out of nine, a moderately negative attitude in each case. She has an implicit moderately negative attitude of four out of nine with respect to casual-business attire as well. This is the selected mix of Jane's implicit and explicit attitudes that come into play as she interacts with Darnell as he comes into her office for his first day at work.

Consistent with the parallel-interventionist process that has been implemented, Jane's implicit and explicit attitudes toward Darnell are activated when he is first perceived – he is Black and dressed in casual business attire. Jane's activated attitudes are her attitudes toward Blacks and toward Darnell's casual business attire. Each of these attitudes then plays into her subsequent behavioral choices in interacting with Darnell. Jane's implicit attitudes towards Darnell's race and attire are moderately negative. Her single explicit attitude toward Blacks is similarly negative. In Jane's determination of how she will interact with Darnell, both her implicit and explicit attitudes come into play, nominally at the same time. Jane's implicit and explicit attitudes are each activated and both are negative. The implicit process more quickly proposes a way forward than the explicit process. While both behavioral responses are negative to the same degree, they may not each lead to exactly the same behavioral outcome. The implicitly- and explicitly-activated are not necessarily identical. In this case, it is the explicit attitude that prevails simply by being an attended process. While the explicit process has taken longer to determine a behavioral path, the timing is such that the implicitly chosen outcome has barely had time to come into play and as Logan (1988) has suggested, the timing challenge is manageable. Jane's explicit process then drives her behavior toward Darnell and she will assign him the less favorable work assignments.

Naheed arrives at the office shortly after Darnell. As with her reaction to Darnell's arrival, Jane's relevant implicit and explicit attitudes are activated on Naheed's arrival. On this occasion, ignoring Naheed's attire, it is only Jane's attitudes toward Naheed as a Muslim that becomes active. Jane's implicit attitude towards Muslims is three out of nine, but her explicit attitude is seven out of nine – she has a stereotypic negative attitude towards Muslims, but there is the potential to explicitly override that negative attitude by the more positive explicit attitude. The implicit attitude that would initiate a negative response is quickly initiated. We can follow the activation of the implicitly and subsequent explicitly activated procedures in Figure 1**.**

Jane has two competing automatic procedures, *my-or-others-intern-auto-neutral* and *my-or-others-auto-firm;* that is, a choice between a more neutral or firmer approach in interacting with Naheed. As seen in Figure 1, there were five cycles of the activation process required for one to reach threshold. Each cycle consisted of four events. The first two events note the arrival of the stimuli for each possible implicit choice and the second pair showing the resulting activation level for each of the competing implicitly-driven procedures. In the yellow highlighted event, shown in complete form at the top of the screen, the activation level of 5.062 for *my-or-others-auto-firm* is seen to have exceeded the threshold of 5.0. This leads to the *proc-selected* event for *my-or-others-auto-firm* shown in the second line of the upper panel sector as occurring

immediately after the selection. The *inhibit-received* events at the bottom of the lower panel complete the selection process among the implicitly activated procedures; that is, the competing selection processes are shut down.

Jane's more positive explicit attitude is activated concurrently with the implicitly activated procedures, but proceeds more slowly in launching its response. The faster operating implicitly-driven response, barely getting underway, is then overridden by the more favorable response driven by her more positive explicit attitude. As above, we claim that Jane's negative implicit response is overridden on the basis of the explicitly generated response having control of consciousness and being in the process of determining behavior. Jane then proceeds to grant Naheed the more interesting task assignments. From an implementation perspective, the more positive explicitly driven response simply has higher priority and the lower priority implicitly activated procedure is set to terminate when suspended.

In the OMAR Task Timeline, Figure 2, we can see how the conflict between Jane's implicitly and explicitly selected procedures plays out over the few seconds as she first encounters Naheed. Each line in the display represents the three second time period over which one of Jane's procedures, as labeled at the left, plays out. The procedure in the first line, *make-selectors-and-drive-selection*, drives the execution of the competing implicit procedures; *process-rule-packets*, four lines from the bottom, drives the concurrent, explicit rule-based process selection. The implicit process executing more quickly initiates *my-or-others-intern-auto-firm*, with priority 10. Within a few hundred milliseconds, the slower explicit rule-based process selects and initiates *my-or-others-intern-controlled-neutral*, with priority 20 and highlighted in yellow in the display. As discussed above, these two outcome processes, the procedures selected to control ongoing behaviors, are in conflict and so defined as such. At the time at which the explicit process is initiated, the conflict is processed resulting in the failure of the implicit process with priority 10 as indicated by its red border. Jane then manages her ongoing behavior via the highlighted, explicitly initiated procedure with priority 20.

In the second stereotype scenario, the original attitude parameters were modified to demonstrate how they can drive alternate behavioral outcomes. In contrast to the first scenario in which Jane had only an implicit attitude toward business-casual attire, in the second scenario, Jane has both an implicit and explicit attitude toward business-casual attire. The newly added explicit attitude is more positive than Jane's implicit attitude of four, the explicit attitude having a value of seven. In the first scenario, it was only Jane's attitude toward race that played in her explicitly-driven behavioral outcome selection. In this scenario, both her attitude toward race and her attitude business-casual attire played in her explicitly-driven behavioral outcome selection. As before, Jane's explicit attitude dominates her implicit attitude and her new more favorable explicit attitude toward business-casual attire is sufficient to activate more favorable behavior toward Darnell—in this scenario he gets a better set of task assignments.

For the second scenario, Jane's attitude toward Muslims was changed a well. In this instance, we simply deleted her positive explicit attitude towards Muslims leaving her with just an implicit attitude toward Muslims with a value of three, a moderately negative attitude. Here we have the interesting case in which there is simply an implicit attitude in play without a competing explicit attitude attempting to drive behavior as well. In this instance, her behavioral selection is driven by her stereotypic attitude toward Muslims in the absence of any mediation through an explicit attitude. In contrast to the first scenario, it is Jane's more negative implicit attitude that leads her to give Naheed less favorable task assignments.

3.5 Modeling Social Interactions among Individuals across Varied Situations

In our modeling we contrast social interactions between strangers and those between individuals who have developed a relationship over time. Further, we examine the role that context plays in mediating those interactions. Consequentially, the situational context, the attributes of the individual perceived, and any relationship that exists among the actors all moderate the activation of attitudes that influence behavioral outcomes, using the mechanisms outlined above. In this section we describe how we incorporate context and episodic memory (remembrance) into our computational model.

Social interactions with strangers must go forward without the advantage of a preexisting history of past interactions. If the newly met individual is a member of an out-group, the subject may harbor a history of prejudices grounded in stereotypes with respect to the out-group (Greenwald, McGhee, & Schwartz, 1998; Tajfel, 1970). Subsequent behaviors toward that individual may be influenced by those prejudices or, as we have seen, they may be moderated by explicitly processed attitudes.

In contrast with stereotype-grounded responses based on attitudes toward others as a group member, behaviors toward an individual with whom one has an ongoing relationship, it is the history of those previous interactions that provide the basis for going forward. We can characterize that relationship as an acquaintanceship with the particular individual, one that persists over an extended time period and across diverse situations. Acquaintanceships subsume the goals and procedures particular to the various situations that the subjects have shared in common. The join of the attribute-values covering the particular situations form the episodic and working memory supporting the subject's goals as the procedures implementing those goals are executed in new encounters. As modeled, we have a process-driven architecture (Edelman, 1987; 1989) with the process-connected memory that spans the shared situations of the acquaintanceship.

We illustrate our progress to date by following Jane and Joe who work together at the office of the previously examined scenarios and later in the day engage again at a Buddhist Center. In their first shared situation, Jane and Joe are colleagues at work. His exact role vis-à-vis Jane is determine by the initialization of the particular scenario. He may work in the same office where, like Jane, he would use his to-do list to determine whether or not he is to be the one to greet the new interns on their arrival. He would also use his to-do list to determine the intern supervisory assignments. Alternatively, he might be a colleague of Jane's stopping by to visit from another office. In the scenario under consideration here, he takes on this later role.

As this scenario proceeds, Jane has welcomed her two new interns and settled them at their respective workplaces when Joe stops by to say hello. He has checked his to-do list and finding that he has an hour before his next meeting, he asks Jane if she would like to go for coffee. They do, in fact, go for coffee and upon returning Jane reminds Joe that she will be at the Buddhist Center that evening greeting arriving visitors and asks if might make it there as well. He says that he will try to make and heads off for his meeting.

Before moving on to the continuation of the scenario, let's briefly stay with Jane and Joe's side trip to get coffee. They stop at the Greek Isle Deli where, of course, there is a broad selection of beverages. While they have nominally gone out for coffee, Jane and Joe each make a choice of beverage. And indeed, this choice is driven by a process very much like that derived from the processing of stereotypic attitudes. Whether from memory or from the menu board, their

attitudes toward beverages are activated, a competition ensues and the first of the possible choices to reach threshold is the one each selects. It is the same implicit process that drives stereotype-driven behaviors that, in this case, drives a simple beverage choice—automaticity in everyday life (Bargh & Chartrand, 1999; Hardin & Banaji, 2013). On this particular outing for "coffee," Jane has an iced tea and Joe has a Coke.

Later that day, Jane goes to the Buddhist Center and after first visiting with the Kenpo (the Center's Abbot), she walks to the front door to greet visitors as they arrive. When Joe arrives, based on their acquaintanceship, she picks up their conversation from earlier in the day related to his possible visit to the Center. She then suggests that he might want to attend a meditation session just getting underway. Jane will greet Joe again at the end of the session as he is ready to depart engaging with him based on their shared experiences. The acquaintanceship that Jane and Joe maintain for one another subsumes the goals, procedures, and remembrances related to their shared office and Buddhist Center experiences and enables them to pursue their actions in the diverse situations in a natural and coherent manner. The framework of the acquaintanceship that each maintains with the other enables them to cross-reference aspects of their shared situations in any particular situation. They adapt to the particular situations, their shared workplace and the Buddhist Center, by activating situation-specific goals and procedures relevant to each situation (Clancey, 1997; Suchman, 1987). In each situation, Jane's and Joe's activated goals and procedures have expectations particularly adapted to that situation. Their interactions, outlined above, were guided in part by this set of expectations. As we are about to see, Jane has a somewhat different but related set of goals and procedures with their attendant expectations that she works from when a new visitor to the Buddhist Center arrives.

In contrast with Jane's interactions with Joe, her approach will be different as she engages Steve who is a visitor that she has not met before—essentially she is forming a new acquaintanceship, albeit, with one likely an in-group member (Greenwald et al., 1998). She starts by leading their introductions to one another and subsequently suggests that he might like to attend an Introduction to Buddhism discussion that is just getting underway. In essence and from a modeling perspective, Jane is establishing a new acquaintanceship. When she greets Steve again after the introductory session that he attended, he is now an acquaintance with the beginnings of a history of interactions. If he comes to the Buddhist Center again or they meet at another location, her interactions will be guided by this newly formed acquaintanceship.

3.6 Scenario Visualization

By their very nature, social interactions among individuals can take a long time to play out across varied situations. The scenarios that we have been building can require the simulator to provide a time frame of anywhere from a few minutes to several days. Working with these extended time frames, the only way to reasonably manage scenario development and evaluation is to run in fast-time. Early-stage debugging and in-depth evaluation of particular, complex model behaviors is more quickly accomplished in a fast-time environment.

On the other hand, the visualization of a scenario with avatars interacting in real-time to act out the roles of a scenario's agents can provide insight into scenario operation and agent behaviors not otherwise available by examining a textual online trace that details agent actions. Omissions or mistakes in the avatars actions become readily apparent through real-time visualization. Real-time operation also provides a venue in which others can rapidly gain insight into and evaluate agent behaviors and take the first steps in assessing their psychological validity. Hence, there are

times when we wish to run in fast-time and others in which real-time operation is more appropriate.

With respect to visualization, the OMAR simulator was first set up to run with the Neverwinter Night (NwN; http://nwn.wikia.com) environment and more recently adapted to run with OpenSim (http://opensimulator.org). Figure 1 provides a screenshot of a scene from the NwN world while Figure 2 provides a screenshot from the OpenSim world. The NwN and OpenSim avatars and scenes were developed by collaborators at Wright State University. The software interfaces between OMAR and NwN and then OpenSim were a joint effort with the developers at Wright State University.

The real-time OMAR-NwN interface is based on a preexisting socket interface in OMAR that has been used in the past when OMAR has been linked to other simulators running in either real-time or fast-time. The signal passing paradigm used to communicate between OMAR and NwN is much like that used within OMAR to coordinate operations among an agent's procedures. While the NwN environment performed reasonably well, current work has been migrated to the OpenSim environment, a software community with a better outlook for continued development and support.

The OMAR-OpenSim interface is based on the http protocol that was available in OpenSim and readily adapted on the OMAR side. OMAR provides an Hyper Text Transfer Protocol (HTTP) server that listens for a URL provided by OpenSim which OMAR then uses to complete the communication loop. OpenSim then selects a scenario to run using an *initialize* command and controls the OMAR simulators operation with *run* and *pause* commands. OMAR then controls OpenSim avatar action by passing a series of avatar commands via the HTTP interface. Most commands run open-loop. For commands of indeterminate duration, such as asking an avatar to walk to a new location, OpenSim provides OMAR with a notification when the command has completed.

Figure 1:Joe and Naheed Conversing in the Neverwinter Night Visualization

Figure 2: Jane Greeting Darnell in the OpenSim Visualization

4.0 CONCLUSION

Nisbett and Wilson (1978) cast a critical light on our tenuous ability to report accurately on our higher order cognitive processes that we are still struggling to adequately address. Single and dual process theories and models derived from experiments that examined implicit processes are but the latest stage in this effort. Yet the several theories and models that trace the process or processes from stimuli to the activation of implicit and explicit attitudes to the final selection and execution of a behavior are rich in their diversity. Sometimes these differences are simply matters of getting the labels exactly right, but often they reflect critical differences in the understanding of and assertions about the underlying processes themselves.

The building of a computational model, whether single process or dual process that accurately reflects the progress to date is reliant on these many research efforts. Yet, all the answers are not in place and more often it is the case that for any central issue there are competing views with a near term resolution being unlikely. And, the model building effort itself demands working code such that observers will recognize agent behaviors as psychologically-grounded and representative of realistic human behavior. In effect, the modeling effort raises new questions at a level of detail that existing theories and models have not yet probed or not adequately addressed. Hopefully, some of these questions will prove worthy of further research effort.

The modeling effort to date captures the impact of stimuli on implicit attitudes and their associative processes. The competition among associative processes that leads to the selection of procedures to be pursued is also well represented. Concurrently, the model processes the stimuli in their impact on explicit attitudes and the process by which a further behavioral outcome is proposed. In the case where explicit attitudes are not activated, the model simply pursues the behavioral outcomes selected by the implicit processes. When explicit attitudes are activated, the explicit process operating under conscious control is presumed to prevail over the concurrent implicit processes by virtue of operating through conscious control.

The current state of model development leaves several challenging issues unaddressed. Mechanizing the impact of implicit processes on concurrent explicit processes will be difficult with a number of questions to be answered. If there is an implicitly proposed behavior, can there sometimes be an explicit critique of the implicitly proposed behavior followed by a revision of the proposed behavior that is then acted upon? To date, the model assumes that the explicit process simply comes up with a “better” idea and it is the one that is pursued by virtue of being the explicit process. Is there a mechanism by which the activated implicit attitudes or their proposed outcome can also influence the explicit behavior selection process? This would be a more subtle process, the modeling of which would certainly profit from additional empirical data. The underlying proposition here is that as we begin to realize the integrated operation of implicit/automatic and explicit/controlled processing in the computational model we might better define this as a *single* process model, a multifaceted interaction of diverse contributors that realistically emulates the path from stimuli to behavioral consequences.

5.0 REFERENCES

Agha, G. A. (1986). *ACTORS: A Model of Concurrent Computation in Distributed Systems.* Cambridge, MA: MIT Press.

Arvind, & Culler, D. E. (1983). *Why Dataflow Architectures.* Computational Structures Group Memo 229-1, Laboratory for Computer Science, Massachusetts Institute of Technology.

Bargh, J. A., & Chartrand, T. L. (1999). The Unbearable Automaticity of being. *American Psychologist, 54,* 462-479.

Bohner, G., & Dickel, N. (2011). Attitudes and Attitude Change. *Annual Review Of Psychology, 62,* 391-417.

Clancey, W. J. (1997). *Situated Cognition: On Human Knowledge and Computer Representation*. New York: Cambridge University Press.

Deutsch, S. E. (1998). Interdisciplinary Foundations for Multiple-Task Human Performance modeling in OMAR. *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society,* Madison, WI.

Deutsch, S. E., & Pew, R. W. (2002). Modeling Human Error in a Real-World Teamwork environment. In *Proceedings of the Twentieth-fourth Annual Meeting of the Cognitive Science Society (pp. 274-279).* Fairfax, VA.

Devine, P. G. (1989). Stereotypes and Prejudice: Their Automatic and Controlled Components. *Journal of Personality and Social Psychology, 56,* 5-18.

Edelman, G. M. (1987). *Neural Darwinism: The Theory of Neuronal Group Selection.* New York: Basic Books.

Edelman, G. M. (1989). *The Remembered Present: A Biological Theory of Consciousness.* New York: Basic Books.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition Advancing the Debate. *Perspectives on Psychological Science*, *8*(3), 223-241.

Fiske, S. T. (1998). Stereotyping, Prejudice, and Discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of Social Psychology* (4th ed., Vol. 2, pp. 357-411). New York: McGraw-Hill.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and Propositional Processes in Evaluation: An Integrative Review of Implicit And Explicit Attitude Change. *Psychological Bulletin, 132,* 692–731.

Gawronski, B., & Bodenhausen, G. V. (2011). The Associative-Propositional Evaluation Model: Theory, Evidence, and Open Questions. *Advances in Experimental Social Psychology, 44,* 59–127.

Gawronski, B., & Creighton, L. A. (2013). Dual Process Theories. In D. E. Carlston (Ed.), *The Oxford Handbook of Social Cognition (*pp. 282-313). New York, NY: Oxford University Press.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*(6), 1464-1480.

Glenberg, A. M. (1997). What Memory is for. *Behavioral and Brain Sciences*, *20,* 1-55.

Hardin, C. D., & Banaji, M. R. (2013). The Nature of Implicit Prejudice: Implications for Personal and Public Policy. *The Behavioral Foundations of Public Policy*, 13-30.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and Deliberative Judgments are Based on Common Principles. *Psychological Review*, *118*, 97–109.

Lieberman, M. D. (2009). What Zombies can't do: A Social Cognitive Approach to the Irreducibility of Reflective Consciousness. In J. Evans & K. Frankish (Eds.), In two minds: *Dual process and beyond* (pp. 293-316). New York: Oxford University Press.

Lieberman, M. D. (2010). Social Cognitive Neuroscience. In S. T. Fiske, D. T. Gilbert, G. Lindzey (Eds.). *Handbook of Social Psychology* (pp 143-193). Hoboken, NJ:John Wiley and Sons, Inc.

Logan, G. D. (1988). Toward an Instance Theory of Automatization. *Psychological Review 95,* 492-527.

Morsella, E. (2005). The Function of Phenomenal States: Supramodular Interaction Theory. *Psychological Review, 112,* 1000 - 1021.

Plant, E. A., & Devine, P. G. (1998). Internal and External Motivation to Respond Without Prejudice. *Journal of Personality and Social Psychology, 75,* 811–832.

Schneider, W., & Shiffrin, R.M. (1977). Controlled and Automatic Human Information Processing: 1. Detection, Search, and Attention. *Psychological Review*, 84, pp1-66.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending, and a General Theory. *Psychological Review, 84,* 127–190.

Steele, G. (1990). *Common Lisp. The Language. Second Edition*. Digital Press.

Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality and Social Psychology Review*, *8*(3), 220-247.

Suchman, L. A. (1987). *Plans and Situated Actions: The Problem of Human-Machine Communication.* Cambridge: Cambridge University Press.

Tajfel, H. (1970). Experiments in Intergroup Discrimination. *Scientific American*, *223*(5), 96-102.

Young, M. J. (1998). The Computational Modeling of Memory: The Role of Long-Term Potentiation and Hebbian Synaptic Modification in Implicit and Schema Memory. *Dissertation Abstracts International, 0126*.

6.0 LIST OF ACRONYMS

ACC	Anterior Cingulate Cortex
HTTP	Hyper Text Transfer Protocol
MIT	Massachusetts Institute of Technology
Omar	Operator Model Architecture
PRISM	Parallel Responses into Skeletal Muscle
SCORE	Simulation Core
SFL	Simple Frame Language
SIT	Supramodular Interaction Theory